**Database Design Review**

CSC 6712 Distributed Storage Systems

**Overview**

You will review the design of an existing distributed storage system to determine the tradeoffs made in its design.  You will give a 15-minute presentation about one system.  Presentations will be done individually (no groups).  See the rubric for specific aspects of the design that you should address.  A list of papers about various distributed systems is provided below.  I chose databases with published papers since it will be easier to find relevant details.

**Database List**

Ghemawat, Sanjay, Howard Gobioff, and Shun-Tak Leung. 2003. “The Google File System.” In *Proceedings of the Nineteenth ACM Symposium on Operating Systems Principles*, 29–43. SOSP ’03. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/945445.945450.

Taft, Rebecca, Irfan Sharif, Andrei Matei, Nathan VanBenschoten, Jordan Lewis, Tobias Grieger, Kai Niemi, et al. 2020. “CockroachDB: The Resilient Geo-Distributed SQL Database.” In *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data*. New York, NY, USA: ACM. https://doi.org/10.1145/3318464.3386134.  
  
Zhou, Jingyu, Meng Xu, Alexander Shraer, Bala Namasivayam, Alex Miller, Evan Tschannen, Steve Atherton, et al. 2023. “FoundationDB: A Distributed Key-Value Store.” *Communications of the ACM* 66 (6): 97–105. https://doi.org/10.1145/3592838.

Weil, Sage A., Scott A. Brandt, Ethan L. Miller, Darrell D. E. Long, and Carlos Maltzahn. 2006. “Ceph: A Scalable, High-Performance Distributed File System.” In *Proceedings of the 7th Symposium on Operating Systems Design and Implementation*, 307–20. OSDI ’06. USA: USENIX Association.

Lakshman, Avinash, and Prashant Malik. 2010. “Cassandra: A Decentralized Structured Storage System.” *ACM SIGOPS Operating Systems Review* 44 (2): 35–40. https://doi.org/10.1145/1773912.1773922.

DeCandia, Giuseppe, Deniz Hastorun, Madan Jampani, Gunavardhan Kakulapati, Avinash Lakshman, Alex Pilchin, Swaminathan Sivasubramanian, Peter Vosshall, and Werner Vogels. 2007. “Dynamo: Amazon’s Highly Available Key-Value Store.” *ACM SIGOPS Operating Systems Review* 41 (6): 205–20. https://doi.org/10.1145/1323293.1294281.  
  
Elhemali, Mostafa, Niall Gallagher, Nick Gordon, Joseph Idziorek, Richard Krog, Colin Lazier, Erben Mo, et al. 2022. “Amazon {DynamoDB}: A Scalable, Predictably Performant, and Fully Managed {NoSQL} Database Service.” In *2022 USENIX Annual Technical Conference (USENIX ATC 22)*, 1037–48. https://www.usenix.org/system/files/atc22-elhemali.pdf.

Dong, Siying, Mark D. Callaghan, Leonidas Galanis, Dhruba Borthakur, T. Savor, and Michael Strum. 2017. “Optimizing Space Amplification in RocksDB.” *Conference on Innovative Data Systems Research*. https://lrita.github.io/images/posts/database/Optimizing-Space-Amplification-in-RocksDB.pdf.

Kreps, Jay, Neha Narkhede, and Jun Rao. 2011. “Kafka: A Distributed Messaging System for Log Processing.” In *Proceedings of the 6th International Workshop on Networking Meets Databases (NetDB ’11)*, 11:1–7. New York, NY: Association of Computing Machinery.

Sumbaly, Roshan, J. Kreps, Lei Gao, Alex Feinberg, Chinmay Soman, and Sam Shah. 2012. “Serving Large-Scale Batch Computed Data with Project Voldemort.” *USENIX Conference on File and Storage Technologies*, February, 18. https://doi.org/10.5555/2208461.2208479.

Ongaro, Diego, and John Ousterhout. 2014. “In Search of an Understandable Consensus Algorithm.” In *2014 USENIX Annual Technical Conference (USENIX ATC 14)*, 305–19. https://www.usenix.org/system/files/conference/atc14/atc14-paper-ongaro.pdf.

Klophaus, Rusty. 2010. “Riak Core: Building Distributed Applications without Shared State.” In *ACM SIGPLAN Commercial Users of Functional Programming*. New York, NY, USA: ACM. https://doi.org/10.1145/1900160.1900176.

Schulze, Robert, Tom Schreiber, Ilya Yatsishin, Ryadh Dahimene, and Alexey Milovidov. 2024. “ClickHouse - Lightning Fast Analytics for Everyone.” *Proceedings of the VLDB Endowment International Conference on Very Large Data Bases* 17 (12): 3731–44. https://doi.org/10.14778/3685800.3685802.

Adams, Colin, Luis Alonso, Benjamin Atkin, John Banning, Sumeer Bhola, Rick Buskens, Ming Chen, et al. 2020. “Monarch: Google’s Planet-Scale in-Memory Time Series Database.” *Proceedings of the VLDB Endowment International Conference on Very Large Data Bases* 13 (12): 3181–94. https://doi.org/10.14778/3181-3194.

Wang, Lei, Guiqiang Ding, Yulong Zhao, Dingzeyu Wu, and Chengrui He. 2017. “Optimization of LevelDB by Separating Key and Value.” In *2017 18th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT)*, 421–28. IEEE. https://doi.org/10.1109/pdcat.2017.00074.

Burrows, M. 2006. “The Chubby Lock Service for Loosely-Coupled Distributed Systems.” *USENIX Symposium on Operating Systems Design and Implementation*, November, 335–50. https://www.usenix.org/event/osdi06/tech/full\_papers/burrows/burrows\_html/.

Corbett, James C., Jeffrey Dean, Michael Epstein, Andrew Fikes, Christopher Frost, J. J. Furman, Sanjay Ghemawat, et al. 2013. “Spanner: Google’s Globally Distributed Database.” *ACM Transactions on Computer Systems*, 8, 31 (3): 1–22. https://doi.org/10.1145/2491245.

Chang, Fay, Jeffrey Dean, Sanjay Ghemawat, Wilson C. Hsieh, Deborah A. Wallach, Mike Burrows, Tushar Chandra, Andrew Fikes, and Robert E. Gruber. 2008. “Bigtable: A Distributed Storage System for Structured Data.” *ACM Transactions on Computer Systems* 26 (2): 1–26. https://doi.org/10.1145/1365815.1365816.

Pelkonen, Tuomas, Scott Franklin, Justin Teller, Paul Cavallaro, Qi Huang, Justin Meza, and Kaushik Veeraraghavan. 2015. “Gorilla: A Fast, Scalable, in-Memory Time Series Database.” *Proceedings of the VLDB Endowment International Conference on Very Large Data Bases* 8 (12): 1816–27. https://doi.org/10.14778/2824032.2824078.